

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:49:45 ; Search time 140.11 Seconds  
(without alignments)  
8.843 Million cell updates/sec

Title: US-09-331-631a-5\_COPY\_76\_144

Perfect score: 381  
Sequence: 1 NQQRDPQOQEQCQRCQRR.....EEQOREDEKYEERMKGDN 69

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/prodata/2/1aa/55\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/58\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/PCTUS5\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	173	45.4	566	1 US-07-955-905A-2	Sequence 2, App11
2	173	45.4	566	1 US-07-955-905A-22	Sequence 22, App1
3	140.5	36.9	587	1 US-07-955-905A-23	Sequence 23, App1
4	103	27.0	1888	1 US-08-056-200-94	Sequence 94, App1
5	103	27.0	1898	2 US-08-800-644-94	Sequence 94, App1
6	102	26.8	1162	2 US-08-728-323A-2	Sequence 2, App11
7	85.5	22.4	740	1 US-08-257-073-5	Sequence 5, App11
8	83.5	21.9	303	1 US-08-109-391A-2	Sequence 2, App11
9	83.5	21.9	303	1 US-08-459-019A-2	Sequence 2, App11
10	83.5	21.9	303	2 US-08-460-428A-2	Sequence 2, App11
11	83.5	21.9	303	3 US-08-458-860A-2	Sequence 2, App11
12	82.5	21.7	432	2 US-08-933-750C-47	Sequence 47, App1
13	82.5	21.7	432	3 US-09-234-613-47	Sequence 47, App1
14	81	21.3	344	5 5210183-2	Patent No. 5210183
15	81	21.3	614	4 PCT-US95-03236-21	Sequence 21, App1
16	81	21.3	683	5 5210183-3	Patent No. 5210183
17	79	20.7	2182	2 US-08-487-826B-16	Sequence 16, App1
18	77	20.2	788	2 US-08-918-914-4	Sequence 4, App11
19	76.5	20.1	83	1 US-08-253-155A-38	Sequence 38, App1
20	76	19.9	186	2 US-08-557-309B-43	Sequence 43, App1
21	76	19.9	186	3 US-08-834-306-43	Sequence 43, App1
22	76	19.9	376	5 5180810-1	Patent No. 5180810
23	75	19.7	312	1 US-08-285-440-4	Sequence 4, App11
24	75	19.7	312	1 US-08-630-349-4	Sequence 4, App11
25	75	19.7	357	1 US-08-612-986-5	Sequence 5, App11
26	75	19.7	357	1 US-08-361-806A-5	Sequence 5, App11
27	75	19.7	357	4 PCT-US95-16806A-5	Sequence 5, App11
28	75	19.7	532	1 US-08-285-440-5	Sequence 5, App11

29	75	19.7	532	1 US-08-630-349-5	Sequence 5, App11
30	75	19.7	558	1 US-08-285-440-6	Sequence 6, App11
31	75	19.7	558	1 US-08-630-349-6	Sequence 6, App11
32	74.5	19.6	288	2 US-08-961-858-6	Sequence 6, App11
33	74.5	19.6	288	3 US-09-089-593-6	Sequence 6, App11
34	74.5	19.6	288	3 US-08-950-923-6	Sequence 6, App11
35	74.5	19.6	450	2 US-08-665-037-2	Sequence 2, App11
36	74.5	19.6	450	2 US-08-666-067-2	Sequence 2, App11
37	74.5	19.6	450	2 US-08-732-870-2	Sequence 2, App11
38	74.5	19.6	905	2 US-08-574-959A-9	Sequence 9, App11
39	74.5	19.6	1135	2 US-08-574-959A-7	Sequence 9, App11
40	74	19.4	514	2 US-08-960-022-14	Sequence 14, App1
41	73.5	19.3	605	1 US-07-955-905A-24	Sequence 24, App1
42	73	19.2	77	1 US-08-253-155A-47	Sequence 47, App1
43	72.5	19.0	571	1 US-07-955-905A-25	Sequence 25, App1
44	72	18.9	360	2 US-08-531-927B-2	Sequence 2, App11
45	71.5	18.8	158	2 US-08-670-186-4	Sequence 4, App11

## ALIGNMENTS

```
RESULT 1
US-07-955-905A-2
; Sequence 2, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ. ID NO.: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-955-905A-2

Query Match          45.4%; Score 173; DB 1; Length 566;
Best Local Similarity 32.6%; Pred. No. 3.7e-10;
Matches 31; Conservative 17; Mismatches 15; Indels 32; Gaps 1;

QY 3 QRPDQOQEQCQRCQRRTEPRHMOICQRCRRYERKRRKQD----- 46
      :|||||:|||||:  || : |||||:|||||:
Db 35 ERDPQOQEQCQRCSEATEERQEQRCRRERKQDQOEELQRYOQCGRQCE 94
QY 47 -----KRYEQOREDEKYEERMK 65
      :|||||:|||||:  |||:|
Db 95 QQOQREQOQCRKCEWYKQERGEHENYHNHK 129

RESULT 2
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

?      COMPUTER:  IBM PC compatible
?      OPERATING SYSTEM:  PC-DOS/MS-DOS
?      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER:  US/07/955,905A
?      FILING DATE:  21-JAN-1993
?      CLASSIFICATION:  435
?      INFORMATION FOR SEQ ID NO:  22:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH:  566 amino acids
?      TYPE:  amino acid
?      TOPOLOGY:  linear
?      MOLECULE TYPE:  protein
?      ORIGINAL SOURCE:
?      ORGANISM:  Theobroma cacao
?      FEATURE:
?      NAME/KEY:  Protein
?      LOCATION:  1..566
?      OTHER INFORMATION:  /note= "67 kD Precursor Protein"
IS-07-955-905A-22

```

Query Match	45.48;	Score 173;	DB 1;	Length 566;
Best Local Similarity	32.68;	Pred. NO. 3.7e-10;		
Matches 31;	Conservative 17;	Mismatches 15;	Indels 32;	Gaps 1

```

Oy      3  QRDPOOYEOCCRCQRETEPRHMOICQRCERYEKEKPKO-----46
          :||:|||||:|:  | | : |:||| | :||:|:|:
Db      35  ERDPROYEOC RCESATTEEREOEOCEQRCERYEYEQROQEFELQROYQCCGRCQ 94

```

```
QY 47 -----KRYEQQREDEEKYEERMK 65
      ::|||:| |
Db 95 QQQGQREQQQCKRKMEQYKEDGERCHENYHNHKK 129
```

RESULT 3  
US-07-955-905A-23

```

?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?
? Sequence 23, Application US/07955505A
? Patent No. 5770433
?
? GENERAL INFORMATION:
?
? APPLICANT:
?
? TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
?
? TITLE OF INVENTION: PRECURSOR
?
? NUMBER OF SEQUENCES: 28
?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
?
? COMPUTER: IBM PC compatible
?
? OPERATING SYSTEM: PC-DOS/MS-DOS
?
? SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
?
? CURRENT APPLICATION DATA:
?
? APPLICATION NUMBER: US/07/955, 905A
?
? FILING DATE: 21-JAN-1993
?
? CLASSIFICATION: 435
?
? INFORMATION FOR SEQ ID NO: 23:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 587 amino acids
?
? Type: amino acid
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: protein
?
? ORIGINAL SOURCE:
?
? ORGANISM: Gossypium hirsutum
?
? FEATURE:
?
? NAME/KEY: Protein
?
? LOCATION: 1..587
?
? OTHER INFORMATION: /note= "Vicilin from G. hirsutum"
?
? OS-07-955-905A-23

```

Query Match	36.9%;	Score 140.5;	DB 1;	Length 587;
Best Local Similarity	28.3%;	Pred. No. 6.4e-07;		
Matches 28;	Conservative 21;	Mismatches 17;	Indels 33;	Gaps 2
1	NR04PQQQIECCQKRCQRETERPHNMICQRCERRERKEKRRKQDKREEDQR-----			

Dd	77	HRPDPKRYECCOEC-KQEROPPOCKPIKPEEOOOOSOKOFEOCHHOIE	134
Qy	55	-----EDKEYEPMKESD	68
Dd	135	QRPEKKOCVAECREPYOENMPREEREAEETEBOE	173

US-08-056-200-94

; Sequence 94, Application US/08056200  
; Patent No. 5616500

ABDICANT: Steinhart Peter M

APPLICANT: Lee, Seung-Chul

```
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
```

APPLICANT: Park, Sang-Chul  
TITLE OF INVENTION: T-10

```

; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same

```

; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobb, Martens, Olson & Bear  
 STREET: 630 Newport Center Drive  
 CITY: 4th South Floor

SIREN: 020 Newport  
CITY: Newport Beach

STATE: CA  
COUNTRY: U.S.A.

ZIP: 92660

MEDIUM TYPE: Floppy

```
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1:25  
CURRENT APPLICATION DATA.

COMMENT AFFILIATION DATA:  
APPLICATION NUMBER: US/08/056,200

;; FILING DATE: 30-APR-1993  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Fredrick Michael F

NAME: FEOLICK, MICHAEL F.  
REGISTRATION NUMBER: 36,799

REFERENCE/DOCKET NUMBER: NIH054.001A  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (714) 760-0404  
TELEFAX: (714) 760-9502

TELEPH: (114) 700-9502  
; INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1898 amino acids

TYPE: amino acid

MOLECULE TYPE: protein

US-08-056-200-94

Query Match 27.0%; Score 103; DB 1; Length 1898;

Best Local Similarity 31.0%; Pred. No. 0.011;  
Matches 22; Conservative 23; Mismatches 20; Indels 6; Gaps 1;

```
QY      56 DEEYEEEMKE 66
          1 : 111 : :
Db      327 RREQQEEERREQ 337
```

RESULT 5  
US-08-800-644-94

Sequence 94, Application US/08800644

; Patent No. 5958752  
; GENERAL INFORMATION:  
;

APPLICANT: Steinert, Peter M.  
APPLICANT: Tee, Sandra-Chul

affidavit, Dec, 2009. Child



APPLICATION NUMBER: US 07/672,183  
FILING DATE: 20-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2570  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 740 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: Internal  
US-08-257-073-5

Query Match 22.4%; Score 85.5; DB 1; Length 740;  
Best Local Similarity 30.4%; Pred. No. 0.23;  
Matches 21; Conservative 18; Mismatches 21; Indels 9; Gaps 2;

QY 3 QRPQOQYEQCQRCQRETE---PRHMOICQRCRCRRYKRRKQKRYEQOREDE 57  
DB 645 QETEEEMKQVEALTKQIEAFVADALAPKN---KEEEKEKEKEKEKEKEKEKE 700

QY 58 EKEERAKE 66  
DB 701 EKEKEKE 709

## RESULT 8

US-08-109-391A-2

Sequence 2, Application US/08109391A

Patent No. 5639876

GENERAL INFORMATION:

APPLICANT: Tripp, Cynthia A.

APPLICANT: Frank, Glenn R.

APPLICANT: Griev, Robert B.

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NOVEL

TITLE OF INVENTION: PARASITIC HELMINTH PROTEINS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross &amp; McIntosh

STREET: 1700 Lincoln St., Suite 3500

CITY: Denver

STATE: CO

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/109,391A

FILING DATE: 19-AUG-1993

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-13

TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/863-9700

TELEFAX: 303/863-0223

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 303 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-109-391A-2

Query Match 21.9%; Score 83.5; DB 1; Length 303;  
Best Local Similarity 23.9%; Pred. No. 0.16;  
Matches 17; Conservative 27; Mismatches 16; Indels 11; Gaps 3;

QY 7 QOQYEQC---QKRCQRETEPR---HMOICQRCRCRRYKRRKQKRYEQOREDEEK 59  
DB 227 EQKTYDQVKEDKERRMQEEERQRLREYEQIQEKLQEEERERQEDERRQKERRMQER 286

QY 60 ---YEERAKE 66  
DB 287 IREQYDEKERQ 297

## RESULT 9

US-08-459-019A-2

Sequence 2, Application US/08459019A

Patent No. 5686080

GENERAL INFORMATION:

APPLICANT: Tripp, Cynthia A.

APPLICANT: Frank, Glenn R.

APPLICANT: Griev, Robert B.

TITLE OF INVENTION: NOVEL PARASITIC HELMINTH P4 PROTEINS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross &amp; McIntosh

STREET: 1700 Lincoln Street, #3500

CITY: Denver

STATE: CO

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,019A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 33,020

REFERENCE/DOCKET NUMBER: 2618-13-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 303 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-459-019A-2

Query Match 21.9%; Score 83.5; DB 1; Length 303;  
Best Local Similarity 23.9%; Pred. No. 0.16;  
Matches 17; Conservative 27; Mismatches 16; Indels 11; Gaps 3;

QY 7 QOQYEQC---QKRCQRETEPR---HMOICQRCRCRRYKRRKQKRYEQOREDEEK 59  
DB 227 EQKTYDQVKEDKERRMQEEERQRLREYEQIQEKLQEEERERQEDERRQKERRMQER 286

QY 60 ---YEERAKE 66  
DB 287 IREQYDEKERQ 297

## RESULT 10

```

US-08-460-428A-2
: Sequence 2, Application US/08460428A
: Patent No. 5912337
:
: GENERAL INFORMATION:
: APPLICANT: Tripp, Cynthia A.
: APPLICANT: Frank, Glenn R.
: APPLICANT: Grieve, Robert B.
: TITLE OF INVENTION: NOVEL PARASTIC HELMINTH
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sheridan Ross P.C.
: STREET: 1700 Lincoln St., Suite 3500
: CITY: Denver
: STATE: CO
: COUNTRY: U.S.A.
: ZIP: 80203
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,428A
: FILING DATE: 02-JUN-1995
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Connell, Gary J.
: REGISTRATION NUMBER: 32,020
: REFERENCE/DOCKET NUMBER: 2618-13-3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 303/863-9700
: TELEFAX: 303/863-0223
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 303 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-460-428A-2
:
: Query Match 21.9%; Score 83.5; DB 2; Length 303;
: Best Local Similarity 23.9%; Fred. No. 0.16;
: Matches 17; Conservative 27; Mismatches 16; Indels 11; Gaps 3
:
QY 7 000YGC-----OKRCORETEPR--IHMOICQRCERREYKREKKRKQKYEEOQREDEEK 59
Db 227 EKKIVQVEKDKERRKQEEBQERLREYERIOEKLQREEBERQROKREKREKREDEER 286
:
QY 60 ---VEERAKE 66
:
Db 287 IRQYDEKERQ 297
:
RESULT 11
US-08-458-860A-2
: Sequence 2, Application US/08458860A
: Patent No. 6100390
:
: GENERAL INFORMATION:
: APPLICANT: Frank, Glenn R.
: APPLICANT: Tripp, Cynthia A.
: APPLICANT: Grieve, Robert B.
: TITLE OF INVENTION: NOVEL PARASTIC HELMINTH
: TITLE OF INVENTION: P22U PROTEINS
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sheridan Ross P.C.
: STREET: 1700 Lincoln St., Suite 3500
: CITY: Denver
: STATE: CO
: COUNTRY: U.S.A.
: ZIP: 80203

```

[illegible]





